ELSEVIER
Linear Algebra and its Applications 335 (2001) 95–100
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

A matrix formulation of Grüss inequality

P.F. Renaud

Department of Mathematics and Statistics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand

Received 22 August 2000; accepted 6 December 2000

Submitted by R.A. Brualdi

Abstract

A matrix trace inequality, which can be regarded as a non-commutative version of Grüss inequality is derived. The proof easily generalises to the more general case of bounded linear operators on a Hilbert space. Other inequalities such as the Kantorovich inequality can be readily derived. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Grüss inequality; Trace inequality

1. Introduction

Suppose that f and g are bounded integrable functions on $[a, b]$ such that

$$\alpha \leqslant f(x) \leqslant \beta, \quad \gamma \leqslant g(x) \leqslant \delta \quad \text{for all } x \in [a, b].$$

Grüss [2] showed that

$$\left|\frac{1}{b-a}\int_a^b f(x)g(x)\mathrm{d}x - \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x \frac{1}{b-a}\int_a^b g(x)\mathrm{d}x\right| \leqslant \frac{(\beta-\alpha)(\delta-\gamma)}{4}. \tag{1.1}$$

In this paper we show that an analogous inequality arises when the functions are replaced by normal matrices and integration by a trace function. If the matrices are not assumed normal, a similar inequality can still be found though with a weaker constant. As will be evident, the proof extends to the more general case of bounded

E-mail address: p.renaud@math.canterbury.ac.nz (P.F. Renaud).

PII: S 0 0 2 4 - 3 7 9 5 (0 1) 0 0 2 7 8 - 6

linear operators on a Hilbert space and Grüss inequality can be recovered from this more general case.

Special cases also lead to the Kantorovich and other inequalities.

For further results and generalisations of Grüss inequality, see [1] and the references therein.

2. A matrix Grüss inequality

Let A and B be $n \times n$ matrices. Write $\|A\|$ for the operator matrix norm of A acting on the inner product space $\mathbb{C}^n$.

Call T (also $n \times n$) a *trace matrix* if T is positive semi-definite with $\mathrm{Tr}(T) = 1$.

Denote by $W(A)$ the numerical range of A, i.e.

$$W(A) = \{\langle Ax, x\rangle \colon \|x\| = 1\},$$

where $\langle\cdot,\cdot\rangle$ is the usual inner product on $\mathbb{C}^n$.

Let

$$w(A) = \sup\{|\langle Ax, x\rangle| \colon \|x\| = 1\}$$

be the numerical radius of A. (For these and other results on numerical ranges, see [3].)

Recall [3, Theorem 1.3-1] that

$$\frac{1}{2}\|A\| \leqslant w(A) \leqslant \|A\|.$$

If A is normal, then [3, 1.4], $w(A) = \|A\|$.

Our first result is as follows.

Theorem 2.1. *Suppose that $W(A)$ and $W(B)$ are contained in disks of radii R and S, respectively. Then, if T is a trace matrix,*

$$|\mathrm{Tr}(TAB) - \mathrm{Tr}(TA)\mathrm{Tr}(TB)| \leqslant 4RS.$$

If A and B are normal, the constant 4 *can be replaced by* 1.

Proof . For ease of notation, we (temporarily) introduce the semi-inner product

$$(A, B) = \mathrm{Tr}(TA^*B).$$

By expanding the left-hand side, we have

$$(A - (I, A)I, B - (I, B)I) = (A, B) - (A, I)(I, B)$$

so that by the Cauchy–Schwarz inequality,

$$\begin{aligned} &|(A, B) - (A, I)(I, B)|^2 \\ &\quad\leqslant (A - (I, A)I, A - (I, A)I)(B - (I, B)I, B - (I, B)I) \\ &\qquad= [(A, A) - |(I, A)|^2][(B, B) - |(I, B)|^2]. \end{aligned} \tag{2.1}$$

Suppose now that $W(A)$ is contained in the disk $\{z: |z-\alpha| \leqslant R\}$.

Substituting $A' = A - \alpha I$, an easy calculation shows that

$$(A, A) - |(I, A)|^2 = (A', A') - |(I, A')|^2 \leqslant (A', A').$$

Furthermore, the numerical range $W(A')$ is now contained in the disk $\{z: |z| \leqslant R\}$. It follows that

$$(A', A') = \operatorname{Tr}(TA'^{*}A') \leqslant \|A'^{*}A'\| = \|A'\|^2 \leqslant 4R^2. \tag{2.2}$$

Similarly, if $W(B)$ is contained in the disk $\{z: |z-\beta| \leqslant S\}$, then

$$(B, B) - |(I, B)|^2 \leqslant 4S^2. \tag{2.3}$$

The result now follows from (2.1)–(2.3) and replacing A by A^*.

That the constant 4 is best possible can be seen from choosing

$$T = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then $w(A) = w(B) = \frac{1}{2}$ so that we may assume that $R = S = \frac{1}{2}$.

Also

$$\operatorname{Tr}(TAB) = 1, \quad \operatorname{Tr}(TA) = \operatorname{Tr}(TB) = 0$$

and the desired inequality holds with equality.

For normal matrices, $w(A) = \|A\|$ so that for such matrices (2.2) now reads

$$(A', A') \leqslant \|A'\|^2 \leqslant R^2$$

and the result follows as before.

In this case the constant 1 is best possible. For if

$$T = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad \text{and} \quad A = B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

then

$$\operatorname{Tr}(TAB) = 1 \quad \text{and} \quad \operatorname{Tr}(TA) = \operatorname{Tr}(TB) = 0.$$

Since $w(A) = w(B) = 1$, $R = S = 1$ and in this case the desired inequality again holds with equality. $\square$

As a corollary to this result, we have the well-known Kantorovich inequality. This can be found in various forms but perhaps the simplest and most useful is the following one (see [4, Theorem 7.4.41]).

Corollary 2.2. *Let A be positive definite with smallest and largest eigenvalues λ_1 and λ_2. If x is a unit vector, then*

$$\langle Ax, x\rangle \langle A^{-1}x, x\rangle \leqslant \frac{(\lambda_1 + \lambda_2)^2}{4\lambda_1\lambda_2}.$$

Proof. The eigenvalues and hence also the numerical range of A lie in a disk centred at $\frac{1}{2}(\lambda_1+\lambda_2)$ and of radius $R=\frac{1}{2}(\lambda_2-\lambda_1)$.

Similarly A^{-1} has its numerical range in a disk of radius

$$S=\frac{\lambda_1^{-1}-\lambda_2^{-1}}{2}.$$

Let T be the projection matrix onto the subspace spanned by x. Clearly T is a trace matrix and for a matrix C, $\mathrm{Tr}(TC)=\langle Cx,x\rangle$.

So putting $B=A^{-1}$ we obtain from Theorem 2.1,

$$|1-\langle Ax,x\rangle\langle A^{-1}x,x\rangle|\leqslant\frac{\lambda_2-\lambda_1}{2}\,\frac{\lambda_1^{-1}-\lambda_2^{-1}}{2}=\frac{(\lambda_2-\lambda_1)^2}{4\lambda_1\lambda_2}$$

so that

$$\langle Ax,x\rangle\langle A^{-1}x,x\rangle\leqslant 1+\frac{(\lambda_2-\lambda_1)^2}{4\lambda_1\lambda_2}=\frac{(\lambda_1+\lambda_2)^2}{4\lambda_1\lambda_2}.\qquad\square$$

3. The more general case

The proof of Theorem 2.1 relies on a semi-inner product defined on the space of matrices via a trace matrix. Finite dimensionality is not required. All that is needed in the proofs is that the operators A and B are bounded linear operators on a Hilbert space and that T is a positive semi-definite bounded linear operator with trace 1. The latter condition is taken to mean that

$$\sum_\alpha\langle Te_\alpha,e_\alpha\rangle=1$$

for any orthonormal basis (e_α). See e.g. [5 p. 175].

We then have the general result.

Theorem 3.1. *Let A, B be bounded linear operators on a Hilbert space H and suppose that their numerical ranges $W(A)$ and $W(B)$ are contained in disks of radii R and S, respectively. If T is a trace operator, then*

$$|\mathrm{Tr}(TAB)-\mathrm{Tr}(TA)\mathrm{Tr}(TB)|\leqslant 4RS.$$

If A, B are normal, the constant 4 *can be replaced by* 1.

The original Grüss inequality (as stated in Section 1) is now easy to recapture.

For, let $H=L_2[a,b]$ and let T be the one-dimensional projection onto the subspace spanned by the constant function 1. Let f,g be the bounded functions in H satisfying

$$\alpha\leqslant f(x)\leqslant\beta,\quad\gamma\leqslant g(x)\leqslant\delta\quad\text{for all } x\in[a,b].$$

Then, if A, B are the Hermitian operators defined via

$$(Ah)(x) = f(x)h(x), \quad (Bh)(x) = g(x)h(x),$$

we have

$$\mathrm{Tr}(TA) = \langle Ae_0, e_0\rangle,$$

where e_0 is the normalised constant function $1/\sqrt{b-a}$.

So

$$\mathrm{Tr}(TA) = \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x,$$

and similarly

$$\mathrm{Tr}(TB) = \frac{1}{b-a}\int_a^b g(x)\mathrm{d}x, \quad \mathrm{Tr}(TAB) = \frac{1}{b-a}\int_a^b f(x)g(x)\mathrm{d}x.$$

Finally note that the spectrum of A is contained in the interval $[\alpha, \beta]$ and hence so too is the numerical range $W(A)$ (see e.g. [3, Theorem 1.4-4]). So $W(A)$ is contained in an interval of radius $R = \frac{1}{2}(\beta - \alpha)$.

Similarly $W(B)$ is contained in an interval of radius $R = \frac{1}{2}(\delta - \gamma)$.

The result now follows from Theorem 3.1, applied to normal operators.

4. An open problem

Theorem 2.1 can be rephrased as follows.

Theorem 4.1. *Suppose that $W(A)$ and $W(B)$ are contained in disks of radii R and S, respectively. Then there is a constant $k(A, B)$ such that for any trace matrix T,*

$$|\mathrm{Tr}(TAB) - \mathrm{Tr}(TA)\mathrm{Tr}(TB)| \leqslant k(A, B)RS.$$

We can assume that $1 \leqslant k(A, B) \leqslant 4$. If A and B are normal, then we can take $k(A, B) = 1$.

It would be interesting to characterise $k(A, B)$ in the case where one of the matrices is not normal. In particular, whether it depends on A and B separately, i.e. whether we can write $k(A, B) = h(A)h(B)$, where $h(A)$, $h(B)$ are suitably defined constants. Is it possible to characterise $h(A)$ in terms of $\|A\|$ and $w(A)$?

References

[1] S.S. Dragomir, Grüss inequality in inner product spaces, Austral. Math. Soc. Gaz. 26 (1999) 66–70.

[2] G. Grüss, Über das maximum des absoluten betrages von $\frac{1}{b-a}\int_a^b f(x)g(x)\mathrm{d}x - \frac{1}{(b-a)^2}\int_a^b f(x)\mathrm{d}x \int_a^b g(x)\mathrm{d}x$, Math. Z. 39 (1934) 215–226.

[3] K.E. Gustafson, D.K.M. Rao, Numerical Range, Springer, New York, 1997.
[4] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, MA, 1988.
[5] J. Weidmann, Linear Operators in Hilbert Spaces, Springer, New York, 1980.